

A MAGÁNHANGZÓK ÉS A MÁSSALHANGZÓK IDŐTARTAMÁNAK EGYENSÚLYA

Kohári Anna

Bevezetés

A beszéd időbeli szerveződését számos kutatás vizsgálta korábban. A vizsgálatok főként a köré szerveződtek, hogy milyen tényezők hogyan befolyásolhatják a beszédhangok időtartamát. A kutatások kimutatták, hogy a beszédhangok időtartamára hatással van többek közt a hangzó minősége, a hangzó környezete, a szótag típusa, a hangsorban elfoglalt helye, a szó hosszúsága, a beszéd tempója, a beszélő személye stb. (vö. Magdics 1966; Kaszai 1982; van Santen 1992; Kovács 2002; Gósy–Beke 2010; Fletcher 2010). Több nyelvben megfigyelték például azt az összefüggést, hogy a magánhangzók hosszabbak zöngés mássalhangzó előtt, mint zöngétlen előtt (holland: Slis–Cohen 1969; francia: Mair–Shadle 1996; amerikai angol: Smith 1997). Ezen összefüggés egyik lehetséges magyarázata, hogy a VC-szekvencián belül az időtartamoknak egyfajta kiegyenlítődése jellemző, azaz a hosszabb zöngétlen zárhangok előtt a magánhangzók időtartama rövidebb, mint az átlagosan kisebb időtartamú zöngés zárhangok előtt (Delattre 1962; Chen 1970). Későbbi kutatások azonban nem támasztották alá, hogy ez a szabályszerűség teljesen univerzális lenne. Az összefüggés ugyanis több nyelvben nem volt egyértelműen mérhető (arab: Port et al. 1980; magyar: Grácz 2012). Habár nem feltételezhetjük, hogy a VC-szekvencián belül minden nyelvben egyfajta időtartambeli kiegyenlítődés jelenne meg, a magánhangzók és a mássalhangzók időtartamaránya más elméletekben is kulcsfontosságú szerepet tölt be, mint például a beszédritmus-elméletekben.

Egy, az 1940-es években felállított beszédritmusmodell szerint minden élő nyelvben a beszéd egyenlő időtartamú részekre tagolódik. Egyéni megfigyelésekre alapozva két eltérő csoportot különítettek el egymástól: a hangsúly-időzítésű és a szótag-időzítésű nyelveket (Pike 1945). A szótag-időzítésű nyelvek esetében a beszédben észlelhető egyenletességet a szótagok időtartamához kapcsolták, és feltételezték, hogy a kiejtett szótagok időtartama közel azonos. Ezzel szemben a hangsúly-időzítésű nyelvekről úgy gondolták, hogy a szabályos, de nem egyenletes ritmusérzetet más alapegység időtartambeli azonossága kelti. Feltételezték, hogy a hangsúlytól hangsúlyig terjedő szakaszok időtartama nagyjából megegyezik (Pike 1945; Abercrombie 1967). A hangsúly-időzítésű nyelvek közé sorolták például az angolt, a hollandot. Szótag-időzítésűnek gondolták a franciát, a spanyolt. A későbbiekben

DOI: 10.15775/Beszkut.2017.119-134

a két kategória mellé egy harmadikat is felállítottak, az ún. moraidőzítésű nyelveket (Ladefoged 1975). A moraidőzítésű nyelvek esetében az egymást követő morák (azaz a szótag súlyát megadó fonológiai egységek) időtartamáról feltételezték, hogy közel azonos. Az akusztikai mérések azonban nem támasztották alá, hogy a különböző beszédritmusosztályokhoz (hangsúly-időzítésű, szótag-időzítésű, moraidőzítésű) tartozó nyelvekben az alapegységek azonos időtartamban valósulnának meg. Egy percepciós kísérlet eredményei viszont összhangban voltak a létrehozott beszédritmus-csoportosítással. Ramus és munkatársai (1999) kísérletében újszülöttek különbséget tudtak tenni hallott felvételek alapján angol (hangsúly-időzítésű nyelv) és japán (moraidőzítésű nyelv) mondatok között, de a holland és az angol mondatok között, azaz két hangsúly-időzítésű nyelv között nem.

Egy későbbi modellben viszont találtak összefüggést a felállított beszédritmusosztályok és akusztikai paraméterek között. A szótagok, hangsúlytól hangsúlyig terjedő szakaszok, illetve a morák vizsgálata helyett ugyanazon tulajdonságait vizsgálták a nyelveknek. A beszéd teljes folyamatát úgy közelítették meg, mint a magánhangzók és a mássalhangzók váltakozása. A magánhangzók és a mássalhangzók időtartamának felhasználásával különböző mérőszámokat hoztak létre, amelyek alapján el tudták helyezni a nyelveket egy skála mentén (Ramus et al. 1999). A mérőszámok szoros összefüggést mutattak a korábban egyéni megfigyelések alapján meghatározott beszédritmusosztályokkal. Az egyik ilyen mérőszám a %V volt, amelyet úgy kaptak meg, hogy összeadták az összes magánhangzó időtartamát, és kiszámolták, hogy a beszéd szünetek nélküli időtartamának hány százaléka a magánhangzók időtartama. Különböző nagyságú nyelvi egységeken is meghatározták ezt az arányt; mondatokra (Ramus et al. 1999) és többmondatos, összefüggő szövegekre is kiszámolták a %V értékét (Grabe–Low 2002). Ez a mérőszám lényegében azt mutatja meg, hogy mennyire kiegyensúlyozott a magánhangzós és a mássalhangzós szakaszok időtartamának aránya. Ha tehát ez a mérőszám 50%-nál kisebb értéket vesz fel, akkor a magánhangzók időtartamának összege kisebb a mássalhangzók időtartamának összegénél. Azt találták, hogy a %V magasabb értékeket vesz fel a hagyományosan szótag-időzítésűekhez sorolt nyelvekben, mint a hangsúly-időzítésűekhez sorolt nyelvekben (Ramus et al. 1999; Grabe–Low 2002). A különböző nyelvek %V-értékei tehát szoros összefüggést mutattak a perceptuális élmények alapján létrehozott beszédritmusosztályokkal.

A beszédritmus-mérőszámok és a hozzájuk kapcsolódó kutatási eredményeket több kritika is érte az elmúlt évszázokban (Arvaniti 2012). Egyrészt felvetődött a kérdés, hogy a felállított mérőszámok – köztük a %V – alapján mennyiben különíthetők el ténylegesen a nyelvek beszédritmusuk szerint. Másrészt felmerült, hogy a vizsgált beszédritmus-mérőszámok kapcsolatban lehetnek a tempóval, és ez befolyásolhatja az eredményeket (Dellwo 2010). Habár ezen elméleti keret megbízhatóságáról, felhasználási lehetősé-

geiről ma is vita folyik (l. White–Mattys 2007; Arvaniti 2012), jelenleg ez az egyetlen olyan ismert módszertan, amellyel a percepción alapuló beszédritmusosztályokhoz akusztikailag mérhető tulajdonságok társíthatók.

A %V mérőszám és a tempó kapcsolatáról meglehetősen kevés mért adattal rendelkezünk. A beszédben a *tempó* terminus alatt sok mindent érthetünk. Az akusztikai méréseken belül figyelembe vehetjük a tempónál a szünetek időtartamát (beszédtempó) vagy figyelmen kívül is hagyhatjuk (artikulációs tempó). A beszédtempó egy időegység alatt elhangzó beszédjelek számát jelenti oly módon, hogy a beszéd idejébe a szünetek ideje is beleszámít. Az artikulációs tempó ezzel szemben azt adja meg, hogy a beszélők adott idő alatt hány hangot, szótagot vagy szót stb. ejtenek ki, ha a szünetek időtartamát nem vesszük figyelembe (Gósy 2004; Fletcher 2010; Gibbon et al. 2014). A %V mérőszám és a tempó kapcsolatának vizsgálatakor ez utóbbit vizsgálták korábban. A %V mérőszám definíciójából adódóan függetlenített a tényleges időtartamok nagyságától, ugyanis azt mutatja meg, hogy egy beszédészlet szünetek nélküli időtartamának hány százaléka a magánhangzós szakaszok időtartama. Ebből az következne, hogy akár gyors vagy lassú valakinek az artikulációs tempója, ugyanúgy lehetne a magánhangzós és a mássalhangzós szakaszok időtartamának aránya pl. 50-50%. A kutatási eredmények azonban ezt nem teljesen támasztják alá. Habár a %V mérőszámot meglehetősen robusztusnak találták a tempóval szemben, mégis találtak valamiféle gyenge kapcsolatot közöttük néhány nyelvben. Dellwo (2010) a francia és cseh nyelvben nem talált összefüggést az artikulációs tempó (CV/s) és a %V között, angol és német nyelvben viszont igen. Ezekben a nyelvekben az artikulációs tempó növekedése tendenciaszerűen együtt járt a %V mérőszám növekedésével. A statisztikai tesztek viszont csak részben igazolták, hogy ez a tendencia jelentős lenne, a szerző éppen ezért elhanyagolhatónak tartja. Az általuk vizsgált anyagban egy-egy beszélő normál, lassabb, még lassabb, gyorsabb és még gyorsabb tempójúnak szánt felolvasott szövegei szerepeltek. Majd a ténylegesen mért adatok alapján öt különböző artikulációs tempójú adatsortot felállítva vizsgálták meg, hogy van-e összefüggés a %V mérőszámmal. Russo és Barry (2008) nyelvenként, illetve nyelvjárásonként három egyenlő részre (lassú, normál, gyors) osztották a különböző beszélők spontán beszédében mért artikulációs tempóadatokat, és azt találták az olasz nyelvben, hogy a %V a tempó növekedésével (hang/s és szótag/s) csökkenő trendet mutat. Igaz, statisztikai próbával nem támasztották alá ezt az állítást. A német nyelvben ellentmondások az eredmények a tekintetben, hogy amikor a hangok számát osztották az időtartammal (hang/s), akkor a tempónövekedés együtt járt a %V kismértékű növekedésével, míg ha a szótagokat vették alapul, akkor csökkent a %V értéke a tempó növekedésével. Hangsúlyozzák, hogy a változás egyik esetben sem jelentős. A mérési eredmények tehát nem igazolják a %V és az artikulációs tempó közvetlen kapcsolatát.

Ez az eredmény azért elgondolkodtató, mert feltételezhető, hogy a lassítást és a gyorsítást a beszédben elsősorban a magánhangzók időtartamának változtatásával érzük el. A beszédhangok esetében ismert, hogy nem egyformán tudnak nyúlni vagy rövidülni. A beszédhangok képzési konfigurációja meghatározza, hogy legkevesebb mennyi idő alatt lehet kiejteni egy hangot, illetve mennyi ideig lehet fenntartani (vö. Klatt 1973; Stevens 1998; Xu 2009; Turk–Shattuck–Hufnagel 2014). Több kísérleti eredmény azt igazolta, hogy a magánhangzók és a mássalhangzók időtartama egyaránt változhat különböző tényezők hatására, de a magánhangzók időtartamában nagyobb változás mutatható ki általánosságban, mint a mássalhangzókéban (pl. Hofhuis et al. 1995). Habár a mássalhangzó képzésmódjától is jelentősen függ, hogy mennyiben nyúlhat meg. A zöngétlen spiránsok elaszticitása például igen nagy. Közlés végén például nagyobb mértékben nyúlhatnak, sőt akár hosszabb időtartamban realizálódhatnak, mint az előttük álló rövid magánhangzó, míg a zöngés vagy zöngétlen felpattanó zárhangokra és a nazálisokra ugyanez nem jellemző (l. Oller 1973; Hofhuis et al. 1995).

A jelen kutatás célja, hogy szisztematikusan megvizsgálja az artikulációs tempó és a %V mérőszám esetleges összefüggéseit a magyar beszédben. Minthogy az artikulációs tempó értékét befolyásolhatja a vizsgált szakasz nagysága, ezért a szakirodalomhoz hasonlóan megkülönböztetünk globális és lokális artikulációs tempót (Pfitzinger 1996, 1999; de Looze 2010; Wagner–Windmann 2011; Gibbon et al. 2014). Amíg a globális tempó a teljes szövegre vagy egy beszélő általános beszédmódjára vonatkozik, addig a lokális artikulációs tempó inkább a produkált szövegen belül megjelenő tempóvariációkat jelenti. Habár nem teljesen egyértelmű, hogy hol húzódik a határ a két kategória között (vö. Fletcher 2010), megkülönböztetésük azért lehet iránymutató, mert egészen más eredményekhez vezethetnek. Ily módon az artikulációs tempó többféleképpen is összefügghet a %V mérőszámmal. Egyrészről feltételezhetjük, hogy a különböző beszélők globális artikulációs tempója, azaz a rájuk jellemző tempó összefüggést mutat azzal, hogy milyen arányban valósítják meg a magánhangzók és a mássalhangzók időtartamának arányát. Másrészről egy beszélőn belül is összefügghet a lassabb vagy gyorsabb tempó a %V mérőszámmal.

Anyag, kísérleti személyek és módszer

Kutatásunkhoz 20 magyar anyanyelvi beszélő (10 nő és 10 férfi) kétperces spontán beszédét, teljes szövegfelolvasását és mondatfelolvasását vizsgáltuk a BEA adatbázisból (Gósy et al. 2012). Életkoruk 20 és 60 év között mozgott (a férfiak átlagéletkora: 36,9 év, a nőké: 37,5 év). Az adatközlők ép hallók és ép beszédűek voltak.

Az artikulációs tempó és a %V mérőszám közti összefüggések kimutatásához kétféleképpen vizsgáltuk az anyagot: teljes egybefüggő szövegeken és tagmondatnyi egységeken. Előbbi esetében a teljes kétperces spontán mono-

lógot és a teljes szövegfelolvasást elemeztük, azaz minden egymást követő tagmondatot megvizsgáltunk. Utóbbi esetében beszédmódonként 20-20 tagmondatot választottunk elemzéseink anyagául. A 10 szótagonál rövidebb tagmondatoknál a mérőszámok nagyon szélsőséges értékeket vehetnek fel. Az általános tapasztalat azt mutatja, hogy az ennél hosszabb tagmondatokra viszont a %V mérőszám elhanyagolható mértékben függ a vizsgált időablak hosszától, ezért az ennél kisebb tagmondatokat kizártuk a vizsgálatokból.

A BEA adatbázisban szereplő szöveg összesen 25 tagmondatra bontható, ebből kiválasztottunk 20 tagmondatot. Mivel több beszélő nem olvasta fel a címet, így a szöveg esetében a tagmondatok számolását a szövegtest első mondatával kezdtük. A tagmondatok kiválasztásának alapját a szótagszám adta a már említett okok miatt. A legkevesebb szótagszámú tagmondatok kizárásra kerültek, így csak a 11 szótagonál hosszabbakat elemeztük. A mondatfelolvasásban csak az egy tagmondatból állókat vettük figyelembe. 20 ilyen mondat szerepelt az adatbázisban, amelyeknek szótagszáma 15 és 27 között mozgott. A spontán beszédből szintén 20 tagmondatot vizsgáltunk, amelyek minimum 10 szótagúak voltak. A tagmondatok elkülönítése spontán beszéd esetében nehezebb, mint felolvasásban. A tagmondathatárok meghatározásában az irányelv az volt, hogy az adott rész egy alany-állítmányi viszonytal rendelkező szövegegység legyen (pl. *Péter pizzát kért az étteremben*). Így szintaktikailag azonos „méretű” (tagmondatnyi) egységeket vizsgáltunk mindegyik korpuszrészéből.

A hanganyagot hangszinten annotáltuk a magyar nyelvre is adaptált MAUS nevű automatikusan szegmentáló szoftverrel (Schiel 1999). Az így kapott hanghatárokat manuálisan ellenőriztük a Praat 5.1 szoftverben (Boersma–Weenink 2015). A magánhangzók időtartamát a formánsszerkezet kezdetéhez és végéhez igazítottuk az oscillogram, a spektrogram és auditív információk segítségével, követve a nemzetközi szakirodalomban szokásos hangelhatárolási kritériumokat (l. Grabe–Low 2002). Amennyiben egy hangnak nem találtam akusztikai lenyomatát a spektrogramon, törlésként kezeltem. A közlést kezdő és a szünet utáni zöngétlen felpattanó zárhangok zárszakaszának idejét úgy határoztuk meg, hogy ugyanazon közlés következő, ugyanolyan képzésű hangzó zárszakaszának idejét hozzáadtuk a zárhang zörej részéhez.

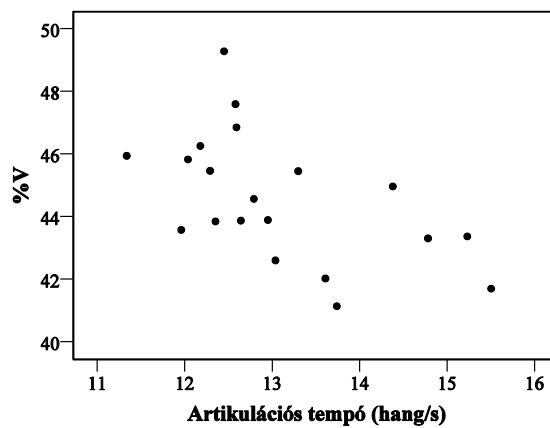
A közlésekben előforduló megakadásokat nem vettük figyelembe az elemzéskor, mivel ezen jelenségek hatása az egyes beszédhangok időtartamára szélsőségesen nagy lehet. Egy-egy kiugróan magas érték olyan mértékben módosíthatná az artikulációs tempó és a %V mérőszám értékét, amely a közlésegység egészére nem jellemző, ezért kihagyásuk mellett döntöttünk. Időtartamukat nem számoltuk bele a beszédhangok artikulációjának tiszta idejébe.

A felszegmentált anyag alapján kétféleképpen határoztuk meg a vizsgálni kívánt paramétereket. Egyrészt kiszámoltuk beszélőnként a teljes szövegfelolvasásra és a teljes spontán beszédre a %V mérőszámot, és összevetettük az ugyanezen anyagra számolt artikulációs tempóval. Másrésztől kiszámoltuk minden beszélő minden tagmondatára mind a %V mérőszámot, mind az artikulációs tempót, és ily módon is összevetettük őket.

A %V definíciója szerint a magánhangzós szakaszok idejének összege osztva az összes beszédhang idejével, és szorozva százszal. A magánhangzós szakaszok alatt az egymást közvetlenül követő magánhangzókat értik (Ramus et al. 1999; Grabe–Low 2002 alapján), de ezen mérőszám esetében nincs jelentősége, hogy magánhangzós szakaszok idejét vagy a magánhangzók idejét adjuk össze, ugyanis ugyanarra az eredményre jutunk. Habár a mérőszám nevében a magánhangzó szerepel, valójában a mássalhangzók és a magánhangzók időtartamára együttesen reagál. Ha ugyanezt a mérőszámot kiszámolnánk a mássalhangzós szakaszokra, akkor az egyenlő lenne a 100%-ból kivont %V értékével. Az artikulációs tempót úgy számítottuk ki, hogy a beszédhangok számát osztottuk az artikuláció tiszta idejére eső időtartammal (Gósy 2004). A statisztikai elemzéseket az SPSS 20.0 program segítségével végeztük. Mivel a statisztikai eljárások elvégzésének előfeltétele az adatok normál eloszlása, ezért a próbák elvégzése előtt normalitásvizsgálatot végeztünk (Shapiro–Wilk-próba alapján). Az adatok nem normál eloszlása miatt nemparametrikus tesztet (Spearman-próba) alkalmaztunk.

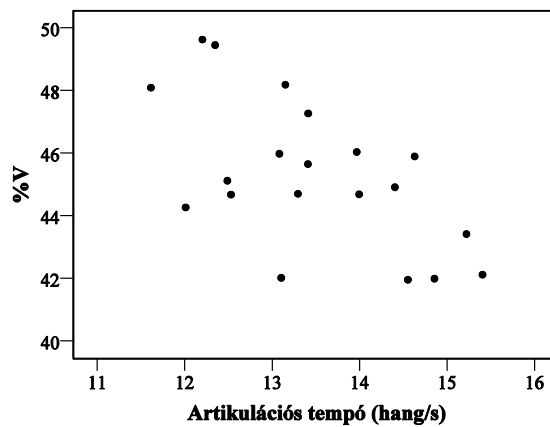
Eredmények

A különböző beszélők eltérő globális artikulációs tempója korrelációt mutatott a %V mérőszámmal. A vizsgálatához 20 beszélő 2 perces monológját és teljes szövegfelolvasását használtuk fel. A beszélők teljes szövegekre és spontán monológokra mért artikulációs tempója közepesen negatívan korrelált az ugyanilyen módon számolt %V mérőszámmal beszédmódtól függetlenül ($\rho = -0,511$, $p = 0,001$). Minél nagyobb volt az adott beszélőre jellemző tempó, annál inkább csökkent a magánhangzós szakaszok (tehát a magánhangzók) időtartamának aránya a teljes beszédidőtartamhoz képest. Megvizsgáltuk külön-külön beszédmódonként, hogy a 20 beszélő globális artikulációs tempója valóban összefügg-e a %V mérőszámmal (1. és 2. ábra). A spontán beszédben statisztikailag is alátámasztható volt a korreláció ($\rho = -0,626$, $p = 0,003$). A szövegfelolvasásban szintén kimutató volt ez az összefüggés ($\rho = -0,512$, $p = 0,021$). A magánhangzók időtartamának százalékos aránya a teljes időtartamhoz képest tehát kisebb a gyorsabban beszélők esetében, és nagyobb a jellemzően lassabb artikulációs tempójú kísérleti személyek esetében.



1. ábra

A %V mérőszám összefüggése az artikulációs tempóval spontán beszédben
(Az ábrán látható pontok egy-egy beszélő 2 perces spontán monológjára számított értékeit mutatják.)



2. ábra

A %V mérőszám összefüggése az artikulációs tempóval szövegfelolvasásban
(Az ábrán látható pontok egy-egy beszélő teljes szövegolvasására számított értékeit mutatják.)

Felmerül a kérdés, hogy vajon csak különböző beszélők globális artikulációs tempója és a %V között van összefüggés, vagy a beszélőkön belüli tem-

póingadozás is kapcsolatban áll a magánhangzók és a mássalhangzók időtartamarányára reagáló mérőszámmal. Ennek megvizsgálásához a beszélők 20-20 tagmondatát elemeztük külön-külön három beszédmódban (szövegfelolvasás, mondatfelolvasás és spontán beszéd). Egy-egy beszélőn belül beszédmódonként összevettük a 20 tagmondatra kapott %V értékeket az ugyanazon tagmondatokra számolt artikulációs tempó értékeivel.

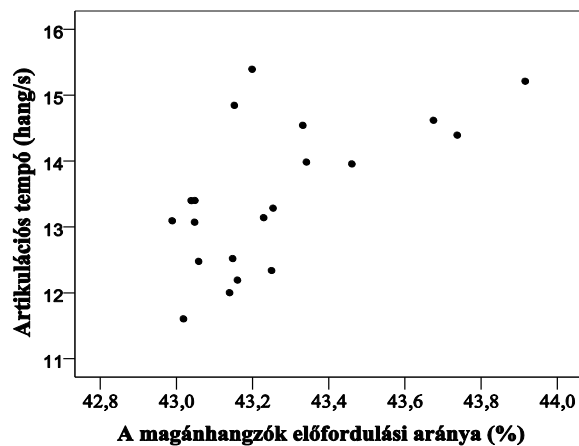
Az eredmények azt mutatták, hogy előfordulnak olyan esetek, amikor egy beszélő lokális artikulációs tempói korrelálnak a %V mérőszámmal, de ez az összefüggés elszórtan, kevésbé szisztematikusan jelent meg, és az esetek többségében nem mutatott összefüggést a statisztikai próba. Nem volt olyan beszélő, akinél mindhárom beszédmódban korrelált volna az artikulációs tempó és a %V mérőszám. Spontán beszédben mindössze két beszélőnél (F4 és F9) volt szignifikáns az összefüggés (F4 esetében: $\rho = -0,554$, $p = 0,010$; F9 esetében: $\rho = -0,511$, $p = 0,021$). Ugyanezen két beszélőnél szövegfelolvasásban is kimutatható volt a korreláció (F4 esetében: $\rho = -0,695$, $p = 0,001$; F9 esetében: $\rho = -0,444$, $p = 0,050$). Mondatfelolvasásban viszont egyik beszélő esetében sem találtunk összefüggést ($p \geq 0,230$). Egyetlen olyan beszélő (N10) volt, akinél az artikulációs tempó és a %V mérőszám mind szövegfelolvasásban ($\rho = -0,629$, $p = 0,003$), mind mondatfelolvasásban ($\rho = -0,468$, $p = 0,038$) összefüggést mutatott. A többi esetben vagy csak a szövegfelolvasásban, vagy csak a mondatfelolvasásban volt korreláció az artikulációs tempó és a %V között a három beszédmód közül. Mondatfelolvasásban összesen csak 4 beszélőnél (F8, N5, N8, N10) volt megtalálható a jelenség ($-0,457 \leq \rho \leq -0,705$, $p \leq 0,043$). Szövegfelolvasásban viszont 8 beszélőnél (F1, F10, N1, N3, N6, N10) korrelált a %V az artikulációs tempóval ($-0,444 \leq \rho \leq -0,695$, $p \leq 0,050$). Az esetek többségében azonban nem sikerült korrelációt kimutatni ($p \geq 0,056$). A %V mérőszám tehát a beszélőkön belül kevésbé függött össze az artikulációs tempóval a vizsgált anyagon, mint a beszélők között.

A %V és az artikulációs tempó összefüggésének háttere

A %V mérőszám – mint láthattuk – korrelációt mutatott a globális artikulációs tempóval. Minél nagyobb volt az adott beszélőre jellemző tempó, annál kisebb volt a magánhangzós szakaszok időtartama a mássalhangzós szakaszok időtartamához képest. Ez az összefüggés többféleképpen is megvalósulhat. Egyrészt előfordulhat, hogy a magánhangzók és a mássalhangzók előfordulásának gyakoriságában térnek el a beszélők. Ha a gyorsabban beszélők több magánhangzót törölnek szemben a lassabb beszélőkkel, akik minden magánhangzót kiejtenek, akkor ennek következtében a magánhangzók időtartamának aránya a teljes beszédidőhöz képest kisebb lenne a gyorsabban beszélőknél. Másrészt az is elképzelhető, hogy a beszélők a mássalhangzók és a magánhangzók időtartamát tendenciózusan másképp valósíthatják meg. A gyorsabban beszélők a magánhangzók időtartamát vagy jobban rövidítik, mint a mássalhangzókét, vagy csak a magánhangzókét rövidítik a mással-

hangzókét nem. Ezen kérdéskör alaposabb tanulmányozásához a teljes szövegfelolvasásokat és spontán monológokat is megvizsgáltuk.

Spontán beszédben a beszédmódból fakadóan a beszélők különböző számú szótagot ejtettek ki. Ezért az előfordulási gyakoriságokat úgy vizsgáltuk meg a %V mérőszám mintájára, hogy kiszámoltuk, hogy az összes beszédhang hány százalékában fordult elő magánhangzó. (A magánhangzók darabszámát elosztottuk a magánhangzók és a mássalhangzók számának összegével, és megszoroztuk százszal.) A magánhangzók darabszámának az összes beszédhanghoz viszonyított aránya nem mutatott korrelációt a %V mérőszámmal spontán beszédben ($\rho = 0,180$, $p = 0,446$). A magasabb, illetve alacsonyabb %V mérőszám tehát nem függött össze azzal, hogy a magánhangzók és a mássalhangzók milyen gyakran fordulnak elő egymáshoz képest egy monológ során. Ez az arány az artikulációs tempóval sem korrelált spontán beszédben ($\rho = 0,128$, $p = 0,591$). Szövegfelolvasásban a magánhangzók darabszámának az összes beszédhanghoz viszonyított aránya szintén nem mutatott összefüggést a %V mérőszámmal ($\rho = -0,089$, $p = 0,710$), az artikulációs tempóval viszont pozitívan korrelált ($\rho = 0,580$, $p = 0,007$). Ez azt jelenti, hogy minél több magánhangzó fordul elő a szövegfelolvasásban a mássalhangzókhoz képest, annál nagyobb volt az artikulációs tempó, tehát gyorsabban beszélt a beszélő (3. ábra). Ez az eltérés azonban nem érintette a %V mérőszám értékeit.

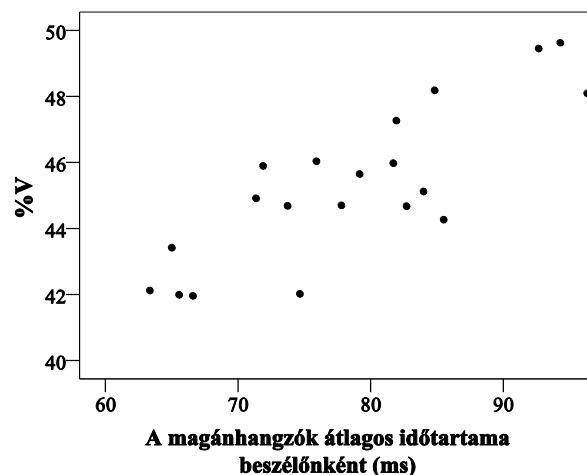


3. ábra

A magánhangzók előfordulási arányának összefüggése az artikulációs tempóval szövegfelolvasásban

A másik tényező, amely a %V és az artikulációs tempó összefüggéseit segíthet jobban megérteni, az a beszédhangok időtartama. A magánhangzók és

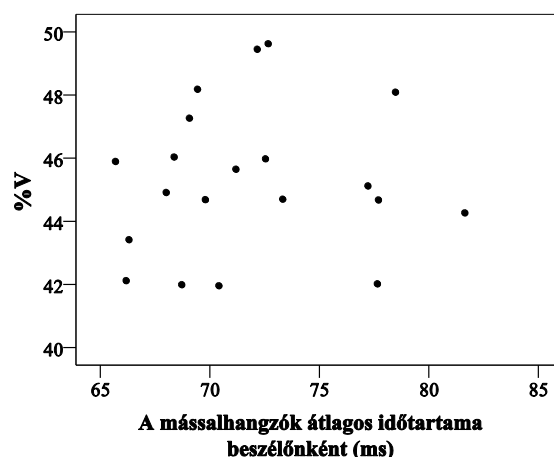
a mássalhangzók időtartama is tendenciózusan eltért a lassabb és a gyorsabb beszélőknél. Beszélőnként külön a kétperces spontán monológ egészére és külön a teljes szövegfelolvasásra megvizsgáltuk a magánhangzók és a mássalhangzók időtartamát. Ezen időtartamok átlagát vetettük össze az adott beszélő globális artikulációs tempójával. A magánhangzók átlagos időtartama erősen és negatívan korrelált az artikulációs tempóval szövegfelolvasásban ($\rho = -0,941, p < 0,001$) és spontán beszédben ($\rho = -0,932, p < 0,001$) is. A mássalhangzók átlagos időtartama szintén erős, negatív korrelációt mutatott az artikulációs tempóval mind spontán beszédben ($\rho = -0,812, p < 0,001$), mind szövegfelolvasásban ($\rho = -0,874, p < 0,001$). Minél gyorsabban beszélt valaki, jellemzően annál inkább kisebb volt a magánhangzók és a mássalhangzók időtartama is a beszédben. A %V mérőszámról viszont tudjuk a jelen cikkben szereplő korábbi eredményeink alapján, hogy minél nagyobb volt egy adott beszélőre jellemző artikulációs tempó, annál kisebb értékeket mutatott, azaz a magánhangzók időtartamaránya a teljes beszédhez képest csökkent. Ez úgy lehetséges, hogy a magánhangzók időtartama nagyobb mértékben csökken, mint a mássalhangzók időtartama a gyorsabb beszélőknél, összevetve a lassabb beszélőkkel. A magánhangzók átlagos időtartama erős, pozitív korrelációt mutatott spontán beszédben ($\rho = 0,808, p < 0,001$) és szövegfelolvasásban is ($\rho = 0,714, p < 0,001$) a %V mérőszámmal. Nem meglepő módon, minél nagyobb volt átlagosan a magánhangzók időtartama, annál nagyobb volt a beszédben a magánhangzók időtartamaránya is (4. ábra).



4. ábra

A magánhangzók átlagos időtartamának összefüggése a %V mérőszámmal szövegfelolvasásban

A mássalhangzók esetében azonban nem igazolódott ennek ellenkezője (5. ábra). A mássalhangzók átlagos időtartama nem függött össze a %V mérőszám értékeivel sem spontán beszédben, ($\rho = 0,177$, $p = 0,457$) sem szöveg-felolvasásban ($\rho = 0,074$, $p = 0,758$).



5. ábra

A mássalhangzók átlagos időtartamának összefüggése a %V mérőszámmal szövegfelolvasásban

Következtetések

A magánhangzós szakaszok időtartamának százalékos aránya a teljes időtartamhoz képest (%V) eredményeink szerint összefüggést mutatott az artikulációs tempóval a magyar beszédben. Korábban viszont nem sikerült kimutatni jelentős összefüggést a %V mérőszám és az artikulációs tempó között más nyelvekben (Russo–Barry 2008; Dellwo 2010). Mérések alapján a %V mérőszámot robusztusnak tekintették az artikulációs tempóval szemben. A magyar beszédben az általunk talált eredmények a klasszikusan szótag-időzítésűnek tartott olaszéhoz hasonlóan (vö. Russo–Barry 2008) azt a trendet mutatták, hogy az artikulációs tempó növekedése jellemzően együtt járt a %V mérőszám csökkenésével. Minél nagyobb az adott beszélőre jellemző tempó, annál inkább kisebb a magánhangzós szakaszok (tehát a magánhangzók) időtartamának aránya a teljes beszédidőtartamhoz képest.

Hozzá kell tennünk, hogy az eredmények alapján a %V és az artikulációs tempó közti összefüggés nem általános érvényű. A különböző beszélők globális artikulációs tempója összefüggést mutatott a %V mérőszám értékeivel, tehát az, hogy ki milyen lassan vagy gyorsan beszél, kapcsolatban van azzal, hogy mi a magánhangzók és a mássalhangzók időtartamának az aránya, de a

beszélőkön belüli tempóingadozás nem mutatott ilyen egyértelmű összefüggést ezzel az aránnyal. Az eredmények azt mutatták, hogy előfordulnak ugyan olyan esetek, amikor egy beszélő lokális tempóértékei korrelálnak a %V mérőszámmal, de ez az összefüggés elszórtan, kevésbé szisztematikusan jelenik meg, és az esetek többségében viszont nem kimutatható. Nem volt olyan beszélő, akinél mindhárom beszédmódban korrelált volna az artikulációs tempó és a %V mérőszám. A %V mérőszám tehát a beszélőkön belül kevésbé függött össze az artikulációs tempóval a vizsgált anyagon. Hangsúlyoznunk kell, hogy beszélőkön belül a %V mérőszám és az artikulációs tempó kapcsolatát számtalan tényező elfedheti. A beszélőkön belüli vizsgálathoz a tagmondatokat egy minimum hosszúság alapján választottuk ki, ily módon az elvégzett elemzések ezen részében – de csakis ezen részében – nem egymást követő tagmondatokat tanulmányoztunk, hanem kisebb időablakokat kiválasztva mértük az analizálni kívánt paramétereket. Az időablakok méretének megválasztása és ezen egységek egymástól való távolsága olyan faktorok, amelyeknek a vizsgált tényezőkre tett hatása jelenleg nem igazán ismert. Eppen ezért elképzelhető, hogy célzott anyaggal és/vagy az egyénenkénti nagyobb mondat vagy tagmondatszámmal sikerülne kimutatni szisztematikus összefüggéseket beszélőkön belül is. A jelen anyag alapján azonban csak annyit állíthatunk, hogy a különböző beszélők globális artikulációs tempója van összefüggésben a magánhangzók és a mássalhangzók időtartamának arányával.

A korábbi szakirodalmi eredményektől való eltérést esetleg magyarázhatja a különböző mérési módszertan. Mint ahogy a kutatási eredményeinkből is látszik, az artikulációs tempó különböző mérési módszerei eltérő eredményekhez vezettek. A vizsgált szakasz nagysága mellett az artikulációs tempónak a különböző mértékegység választása (CV/s, szótag, hang/s) is befolyásolhatta az eredményeket. Ugyanakkor az is valószínűsíthető, hogy a nyelvek között is található különbség az időzítési szabályszerűségben. Eppen ezért a továbbiakban érdemes lenne az artikulációs tempó és a %V mérőszám kapcsolatát más, lehetőleg különböző beszédritmusosztályba sorolt nyelv vizsgálatára is kiterjeszteni oly módon, hogy többféle mértékegységet és szakasznagyságot veszünk figyelembe.

Az eredmények alapján felmerülhet a kérdés, hogy a mássalhangzók és a magánhangzók időtartamának aránya hogyan tolódhat el a gyorsabb vagy a lassabb artikuláció hatására, mi okozza az arány eltéréseit a különböző tempójú beszélőknél. A magánhangzó időtartamok ilyen arányú csökkenésének kézenfekvő a magyarázata lehetne, hogy a gyorsabb artikulációs tempó esetében a beszélők hajlamosabbak törölni a magánhangzókat, és ez mutatkozna meg az időtartamarányokban is. Szövegfelolvasásban minél nagyobb volt az artikulációs tempó, tehát gyorsabban beszélt a beszélő, annál több magánhangzó fordult elő a mássalhangzókhoz képest. Spontán beszédben viszont ez a tendencia nem érvényesült. Továbbá a %V mérőszám nem függött össze

azzal, hogy a magánhangzók és a mássalhangzók milyen gyakran fordulnak elő egymáshoz képest. Így arra következtethetünk, hogy a törlés, illetve a magánhangzók előfordulásának az aránya a mássalhangzókhoz képest kevésbé játszik szerepet a %V mérőszám és az artikulációs tempó közt feltárt összefüggésben.

A másik lehetséges kézenfekvő magyarázat, hogy a gyorsabb, illetve lassabb beszélők a mássalhangzók és a magánhangzók időtartamát tendenciózusan másképp valószínűsíthetik meg. Az eredmények nem meglepő módon azt mutatták, hogy minél gyorsabban beszélt valaki, jellemzően annál inkább kisebb volt a magánhangzók és a mássalhangzók időtartama is a beszédben. A gyorsabb beszélők tehát rövidebben ejtik ugyan a magánhangzókat és a mássalhangzókat is, de tendenciózusan jobban rövidítik a magánhangzókat a mássalhangzóknál. Így jöhet létre a magánhangzók és mássalhangzók időtartamarányának eltolódása gyorsabb vagy lassabb beszédben. Az eredmények azt is feltárták, hogy minél nagyobb volt átlagosan a magánhangzók időtartama, annál inkább nagyobb volt a beszédben a magánhangzók időtartamaránya is. A mássalhangzók átlagos időtartama viszont nem függött össze a %V mérőszám értékeivel sem spontán beszédben, sem szövegfelolvasásban. A magánhangzók időtartama flexibilisebben változtatható, mint a legtöbb képzésmódú mássalhangzóé (vö. Hofhuis et al. 1995), így nem meglepő, hogy a különböző beszélők következetesebben tudják variálni a magánhangzók időtartamát, mint a mássalhangzókéét. Elsősorban tehát a magánhangzók rövidebb vagy hosszabb ejtése áll a háttérben annak, hogy a gyorsabb beszélőknél a magánhangzó időtartamaránya a mássalhangzókéhoz képest kisebb, mint a lassabb beszélőknél.

Összességében tehát elmondható, hogy az egyik ismert beszédritmus-elmélet központi szerepet játszó mérőszáma (%V) korrelál az artikulációs tempóval. Korábban ugyanezen keret több mérőszáma esetében is kimutatták, hogy korrelál az artikulációs tempóval. Annak érdekében, hogy közvetlenül csak a beszédritmust mérhessék olyan variációit hozták létre a mérőszámoknak, amelyek függetlennek tekinthetők a tempótól (Grabe–Low 2002; Dellwo 2010). Kérdés azonban, hogy vajon lehet-e teljesen külön kezelni az ebben a keretben meghatározott beszédritmus tulajdonságokat és az artikulációs tempót. Percepciók kísérletek is arra utalnak, hogy a hallgatók hajlamosak a klasszikusan különböző beszédritmusosztályba sorolt nyelveket a beszéd tempója alapján megkülönböztetni, mintsem a beszédritmus-mérőszámokkal mérhető tulajdonságok alapján (White et al. 2012). Kutatási eredményeink – más korábbi vizsgálatokhoz hasonlóan (vö. Arvaniti 2012; White et al. 2012; Nolan 2014) – arra világítanak rá, hogy az általunk ismert mérőszámokkal mérhető beszédritmus és a tempó kapcsolata korántsem tekinthető teljesen tisztázottnak.

Irodalom

- Abercrombie, David 1967. *Elements of general phonetics*. Edinburgh University Press, Edinburgh.
- Arvaniti, Amalia 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40. 351–373.
- Boersma, Paul – Weenink, David 2015. *Praat: Doing phonetics by computer*. 5.4.22. verzió. <http://www.praat.org/> (A letöltés ideje: 2015. október 8.)
- Bóna Judit 2009. *A gyors beszéd. Produkciós és percepció sajátosságok*. MTA Könyvtára–Lexica Kiadó, Budapest.
- Chen, Matthew 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22/3. 129–159.
- Delattre, Pierre 1962. Some factors of vowel duration and their cross-linguistic validity. *Journal of the Acoustical Society of America* 34/8. 1141–1143.
- Dellwo, Volker 2010. *Influences of speech rate on the acoustic correlates of speech rhythm: An experimental phonetic study based on acoustic and perceptual evidence*. Dissertation. Universität Bonn, Bonn.
- Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In Hardcastle, William J. – Laver, John – Gibbon, Fiona E. (eds.): *The handbook of phonetic sciences*. 2nd edition. Wiley-Blackwell, Oxford. 521–602.
- Gibbon, Dafydd – Klessa, Klessa – Bachan, Jolanta 2014. Duration and speed of speech events: A selection of methods. *Lingua Posnaniensis* 56/1. 59–83.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácsi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Grabe, Esther – Low, Ee Ling 2002. Durational variability in speech and the Rhythm Class Hypothesis. *Papers in Laboratory Phonology* 7. 515–546.
- Grácsi Tekla Etelka 2012. *Zöreghangok akusztikai fonetikai vizsgálata a zöngességi oppozíció függvényében*. PhD-értekezés. ELTE, Budapest.
- Hofhuis, E. M. F. J. – Gussenhoven, Carlos – Rietveld, Toni 1995. Final lengthening at prosodic boundaries in Dutch. In Elenius, Kjell – Branderud, Peter (eds.): *Proceedings of 13th International Congress of Phonetic Sciences*. Stockholm. 154–157.
- Kassai Ilona 1982. A magyar beszédhangok időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest. 13–23.
- Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54. 1102–1104.
- Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.
- Ladefoged, Peter 1975. *A course in phonetics*. Harcourt Brace Jovanovich, New York.
- de Looze, Celine 2010. *ADoReVA and ADoTeVA, two PRAAT plugins for the automatic detection of register and tempo variations*. http://celinedelooze.com/2010/DeLooze_Submission_SPSASSD.pdf (A letöltés ideje: 2016. december 1.)
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.

- Mair, Sheila J. – Shadle, Christine H. 1996. The voiced/voiceless distinction in fricatives: EPG, acoustic and aerodynamic data. In: *Proceedings of Institute of Acoustics* 18. 163–169.
- Nolan, Francis – Jeon, Hae-Sung 2014. Speech rhythm: a metaphor? *Philosophical Transactions of Royal Society B: Biological Sciences* 369/1658. 20130396.
- Oller, D. Kimbrough 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54/5. 1235–1247.
- Pfützinger, Hartmut R. 1996. Two approaches to speech rate estimation. In McCormack, Paul – Russel, Alison (eds.): *Proceedings of the 6th Australian International Conference on Speech Science and Technology* (SST '96). Adelaide, Australia. 421–426.
- Pfützinger, Hartmut R. 1999. Local speech rate perception in German speech. In Ohala, J. John – Hasegawa, Yoko – Ohala, Manjari – Granville, Daniel – Bailey, Ashlee C. (eds.): *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco. 893–896.
- Pike, Kenneth 1945. *The intonation of American English*. University of Michigan Press, Ann-Arbor.
- Port, Robert F. – Al-Ani, Salman – Maeda, Shosaku 1980. Temporal compensation and universal phonetics. *Phonetica* 37/4. 235–252.
- Russo, Michela, – Barry, William J. 2008. Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. In: *Proceedings of Speech Prosody* 4/8. 419–422.
- van Santen, Jan 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.
- Schiel, Florian 1999. Automatic phonetic transcription of non-prompted speech. In Ohala, J. John – Hasegawa, Yoko – Ohala, Manjari – Granville, Daniel – Bailey, Ashlee C. (eds.): *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco. 607–610.
- Slis, Iman Hans – Cohen, Antonie 1969. On the complex regulating the voiced-voiceless distinction. *Language and Speech* 12. 80–102.
- Smith, Caroline L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25/4. 471–500.
- Stevens, Kenneth N. 1998. *Acoustic Phonetics*. The MIT Press, Cambridge, Massachusetts–London.
- Turk, Alice E. – Shattuck-Hufnagel, Stefanie 2014. Timing in talking: What is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society B: Biological Sciences* 369/1658. 20130395.
- Xu, Yi 2009. Timing and coordination in tone and intonation. An articulatory-functional perspective. *Lingua* 119/6. 906–927.
- Wagner, Petra – Windmann, Andreas 2011. The shrinking effects on speech tempo perception. In Lee, Wai-Sum – Zee, Eric (eds.): *Proceedings of the 17th International Conference on Phonetic Sciences*. Hong Kong. 2082–2085.
- White, Laurence – Mattys, Sven L. 2007. Calibrating rhythm: first language and second language studies. *Journal of Phonetics* 35. 501–522.

The balance between the duration of vowels and consonants

The ratio of vowel and consonant durations plays a key role in the speech rhythm model set up by Ramus et al. (1999) and its follow-up variations. In this study, we analysed the relationship between articulation rate and the speech rhythm metric %V in different speech styles. We found statistically significant correlation between the two parameters, implying that speakers exhibiting higher articulation rates tend to have lower values of relative vowel duration %V. The results highlight the complexity of connections between speech rhythm and articulation rate.